# Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Geleris J, Sun Y, Platt J, et al. Observational study of hydroxychloroquine in hospitalized patients with Covid-19. N Engl J Med. DOI: 10.1056/NEJMoa2012410

Supplementary Appendix for "Observational study of hydroxychloroquine in hospitalized patients with COVID-19"

Table of contents


| | Page |
|---|---|
| **Supplementary Methods** | 2 |
| Data extraction | |
| Propensity score modeling | |
| Time-to-event modeling | |
| **Figure S1**: Distribution of the estimated propensity score for receiving hydroxychloroquine, among patients who did and did not actually receive the treatment | 4 |
| **Figure S2:** Standardized mean differences in the unmatched and matched sample | 5 |
| **Figure S3**: Histogram of time (in days) to first dose of hydroxychloroquine for those patients that received hydroxychloroquine before the composite endpoint | 6 |
| **Figure S4:** Distribution of inverse probability score weights in the HCQ and no HCQ groups. The weights were truncated by resetting the value of weights greater (lower) than percentile 99 (1) to the value of percentile 99 (1) | 7 |
| **Table S1:** Final disposition of patients included in the analysis at the end of the study follow up period | 8 |
| **Table S2:** Odds ratios (95% CIs) of receiving hydroxychloroquine treatment for all variables included in the propensity score model | 9 |
| **Table S3:** Sensitivity analyses showing hazard ratio for primary outcome using different baselines than the main results and, separately, only considering patients who received treatment prior to baseline time | 10 |
| **Table S4:** Hazard ratios (95% CIs) of all variables included in the inverse probability weighted Cox model | 11 |

**Supplementary Methods**

Data extraction

The Columbia clinical data warehouse[1] comprises over 30 years of data on over six million patients from the NewYork-Presbyterian / Columbia University Irving Medical Center, collected from electronic health records over time, currently from Epic Systems (Verona, WI). The data include all outpatient and inpatient demographics, visit information, diagnoses, procedures, medications, vital signs, care provider notes, orders and prescriptions, laboratory results, radiology reports, and numerous other ancillary reports. Laboratory and ancillary data are fed directly to the warehouse from the source computing systems and serve as the gold standard for data quality reviews for clinical trials. The Observational Health Data Sciences and Informatics initiative data quality tool set called Achilles Heel[2] includes an extensive knowledge base of data consistency checks used to verify the quality of the Columbia warehouse. Data are requested from the warehouse via a formal specification that is approved by an institutional committee and executed by an analyst.

Extracted vital signs on presentation included heart rate, temperature, respiratory rate, blood pressure and peripheral oxygen saturation. Evaluated laboratory tests included creatinine, D-Dimer, ferritin, C-reactive protein, lactate dehydrogenase, and procalcitonin. We have included additional laboratory tests in this supplement that were not included in the propensity score. Initial PaO2/FiO2 ratio was calculated by estimating the PaO2 from the first recorded oxygen saturation available and then dividing by the estimated FiO2 for that oxygen delivery method (0.21 for room air, 0.21 + (oxygen flow rate * 0.03) for nasal cannula, 0.80 for non-rebreather mask[3,4] or the recorded FiO2 for non-invasive or invasive ventilation). Diagnostic categories were created by grouping ICD-10-CM diagnosis codes using the Clinical Classifications Software (CCS) by the Healthcare Cost and Utilization Project and joining related categories such as Cancer that has many CCS codes. These diagnosis categories included chronic lung disease (including asthma, chronic obstructive pulmonary disease [COPD] and chronic bronchitis), cancer, chronic kidney disease, essential hypertension, hypertension with complications, diabetes without complications, diabetes with complications, pulmonary heart disease and HIV, organ transplant or other immune suppression. Medications evaluated included angiotensin converting enzyme inhibitors and angiotensin-II receptor blockers, statins, steroids, oral anti-coagulation, antibiotics (with the exclusion of azithromycin as it is categorized separately), other potential treatments for COVID-19 (tocilizumab and remdesivir), azithromycin and hydroxychloroquine. We classified patients for each medication depending on whether they received the medication at any point during their hospital stay. Because of the dramatic increase in the need for ICU beds, multiple units that are not typically ICUs were converted to ICUs at various times as patient volumes increased, so identification of ICU patients was not possible. Use of intubation and invasive mechanical ventilation was detected using a combination of ventilator orders, ventilator flowsheets and intubation notes. In clinical practice, the use of invasive mechanical ventilation was left to physician discretion.

Propensity score modeling

We fitted a logistic regression model of HCQ regressed on other baseline covariates and obtain the predicted probability of HCQ. The predicted probabilities were used to calculate the stabilized IPW weight in the time-to-event analysis

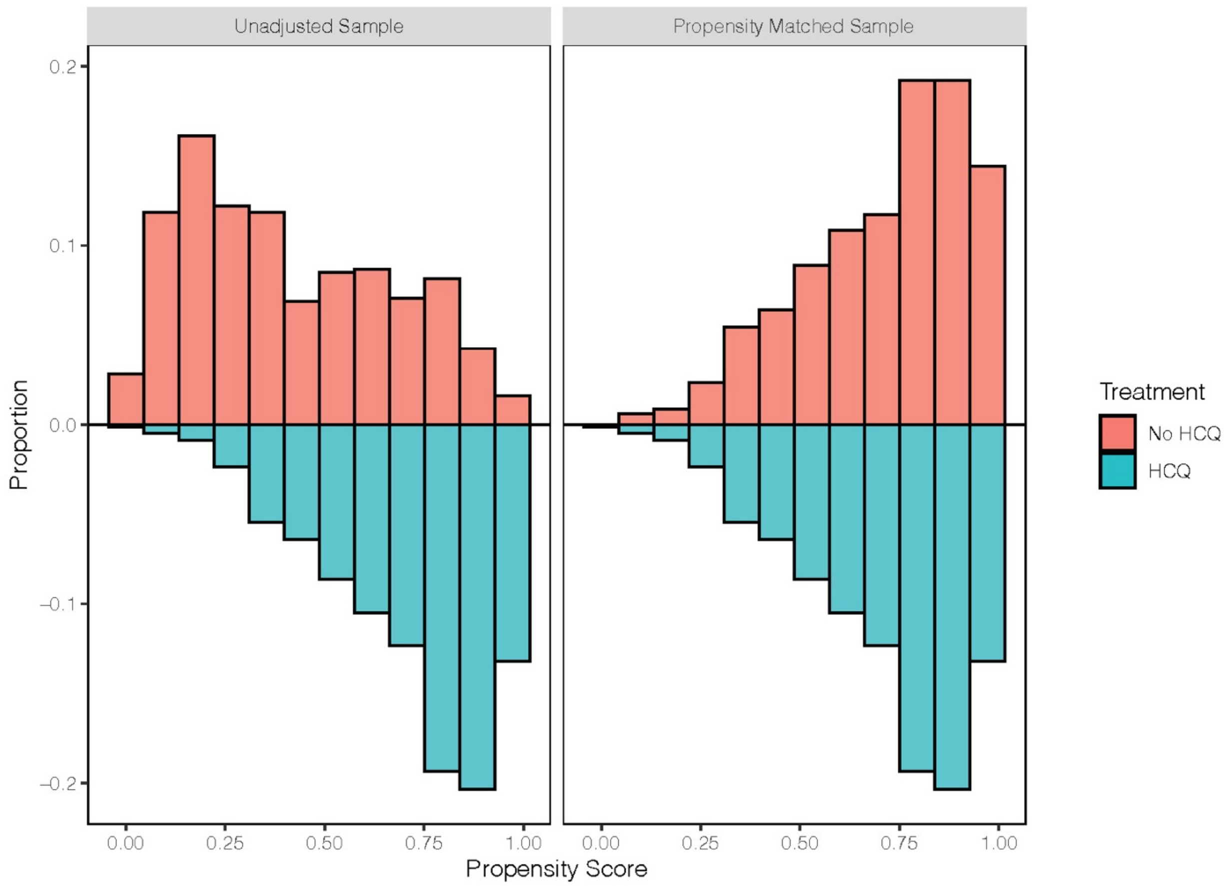
Time-to-event modeling

Cox models: The multivariable Cox models were estimated by adjusting for baseline covariates (or using strata as appropriate). In IPW analysis, Cox models were estimated with case weights based on the stabilize weights. Since multiple imputation was applied to fill in missing data, the final standard error was obtained using Rubin's rule based on the robust variance estimator in Cox regression.

We applied the test for proportionality assumption based on the Schoenfeld residuals. To fix the potential nonproportional hazards, we used strata based on sex, asthma/COPD, and BMI (>30 vs <=30).

IPW Kaplan-Meier curves: We applied nonparametric bootstrap to construct confidence intervals for the IPW Kaplan-Meier curves. For each bootstrap sample obtained from sampling the original 1376 subjects with replacement, we imputed ten datasets, fitted the propensity score model and estimated the curves on each imputed dataset, and then averaged the ten estimated values to obtain the pooled estimate for the bootstrap sample. The confidence intervals are estimated based on 200 bootstrap samples using the normal approximation.

Figure S1: Distribution of the estimated propensity score for receiving hydroxychloroquine, among patients who did and did not actually receive the treatment




On the left, histograms of propensity scores for the unadjusted populations who were treated with hydroxychloroquine and were not treated with hydroxychloroquine. On the right, histograms of the propensity matched samples. Generated using the first imputed dataset. The other imputed datasets are similar and thus omitted.

Figure S2. Standardized mean differences in the unmatched and matched sample

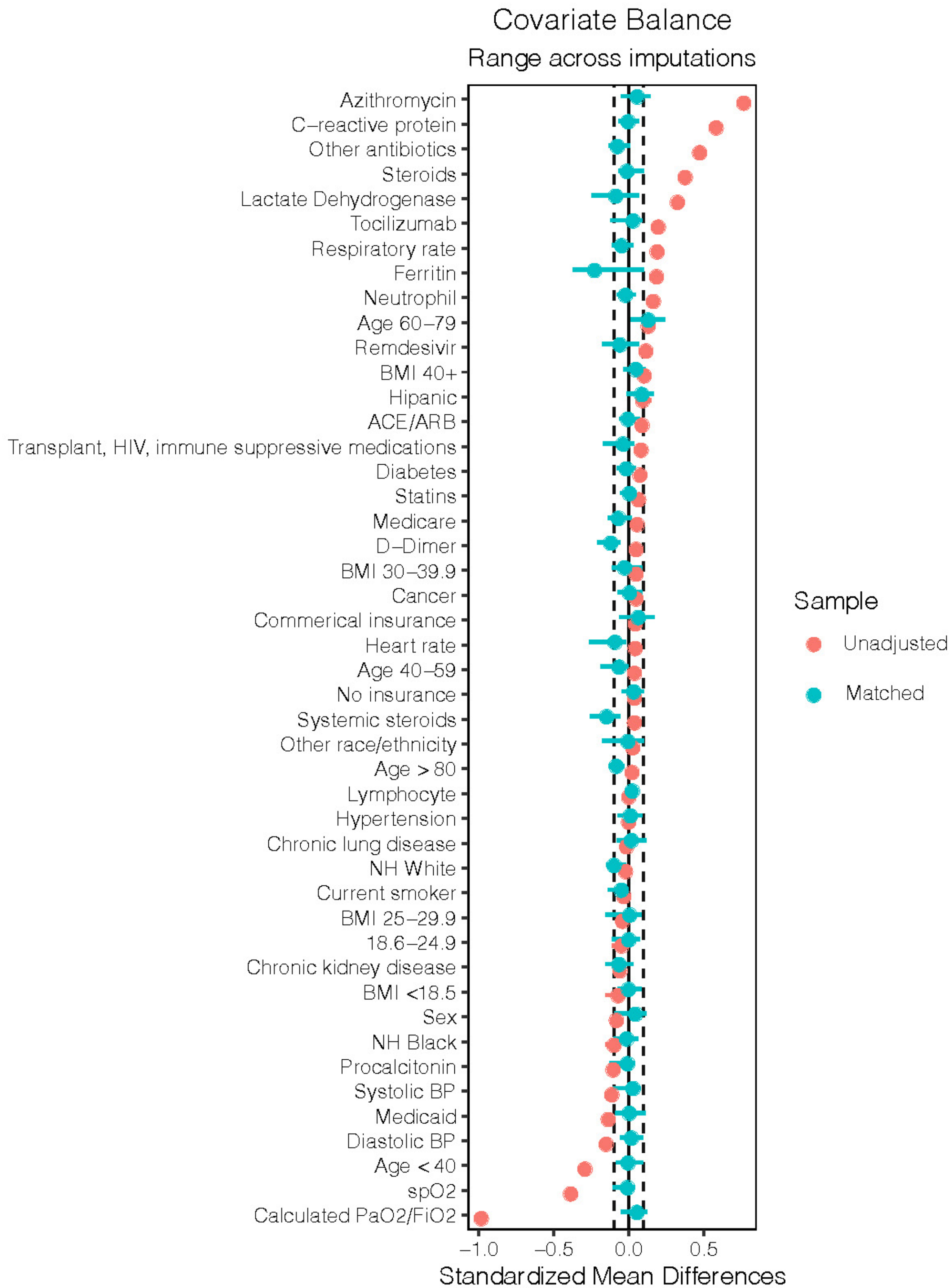


Figure S3. Histogram of time (in days) from arrival to the emergency room to first dose of hydroxychloroquine for those patients that received hydroxychloroquine before the composite endpoint. Study baseline is defined as day 1 on this figure.

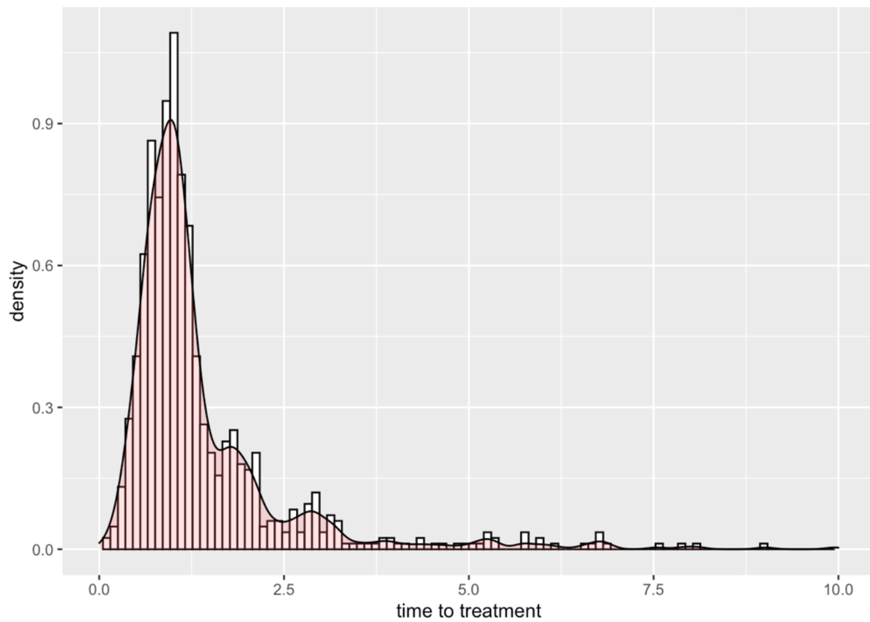


Figure S4. Distribution of inverse probability score weights in the HCQ and no HCQ groups. The weights were truncated by resetting the value of weights greater (lower) than percentile 99 (1) to the value of percentile 99 (1).

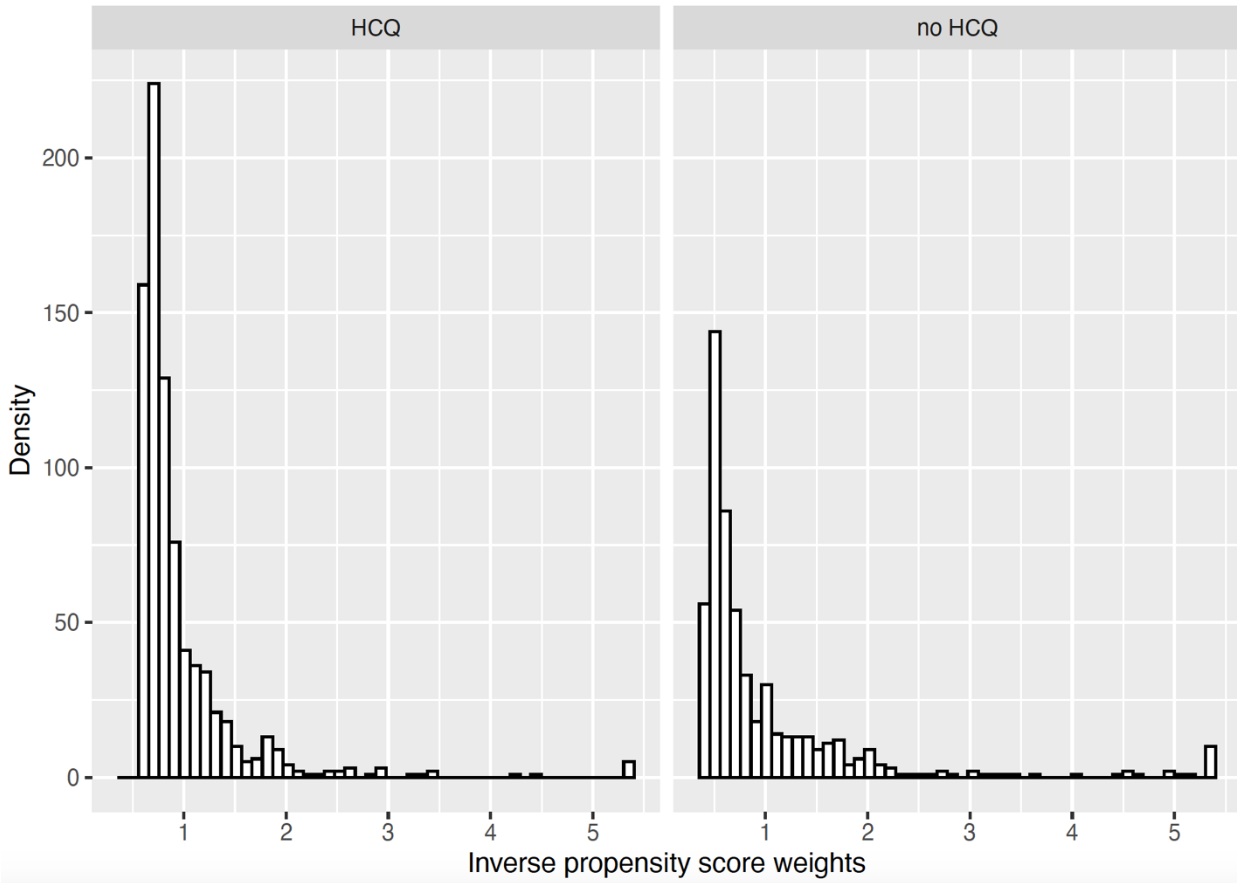


Table S1. Final disposition of patients included in the analysis at the end of the study follow up period. (no. of patients)

| | Hydroxychloroquine | No hydroxychloroquine |
|---|---|---|
| Primary Outcome | 262 | 84 |
| Death | 157 | 75 |
| Intubated | 154 | 26 |
| Intubated then Death | 49 | 17 |
| Still hospitalized and not intubated | 13 | 11 |
| Discharged Alive | 552 | 473 |
| Still hospitalized | 102 | 17 |

**Please note that patients may be in more than one category (e.g. Intubated and Discharged Alive)*

Table S2. Odds ratios (95% CIs) of receiving hydroxychloroquine treatment for all variables included in the propensity score model.

| | (N = 1376) |
|---|---|
| Intercept | 0.12 (0.01-2.04) |
| Age (ref=<40) | |
| 40-59 | 1.18 (0.76-1.83) |
| 60-79 | 1.28 (0.79-2.08) |
| >80 | 1.39 (0.77-2.53) |
| Female sex | 1.13 (0.87-1.47) |
| Race/ethnicity (ref=NH White) | |
| NH Black | 0.68 (0.43-1.07) |
| Hispanic | 1.03 (0.70-1.53) |
| Other | 1.45 (0.34-6.23) |
| BMI (ref=25-29.9) | |
| <18.5 | 0.80 (0.56-1.13) |
| 18.5-24.9 | 0.88 (0.65-1.18) |
| 30-39.9 | 1.05 (0.48-2.30) |
| 40+ | 1.35 (0.80-2.31) |
| Insurance (ref=Medicaid) | |
| Medicare | 0.98 (0.67-1.44) |
| No insurance | 1.20 (0.74-1.94) |
| Commercial | 1.55 (1.05-2.29) |
| Current Smoker | 0.90 (0.60-1.36) |
| Past Diagnoses | |
| Asthma, COPD, or bronchiectasis | 0.81 (0.58-1.14) |
| Essential hypertension | 0.93 (0.69-1.26) |
| Cancer | 1.06 (0.72-1.56) |
| Chronic kidney disease | 0.76 (0.53-1.09) |
| Transplant, HIV, or immune suppressive medications | 2.07 (1.05-4.09) |
| Current Medications | |
| Statins | 1.18 (0.87-1.59) |
| ACE/ARB | 1.02 (0.76-1.38) |
| Steroids | 2.56 (1.75-3.75) |
| DOAC or Coumadin | 1.06 (0.68-1.65) |
| Azithromycin | 2.72 (2.04-3.62) |
| Other Antibiotics* | 1.63 (1.25-2.12) |
| Tocilizumab | 1.16 (0.55-2.47) |
| Remdesivir | 1.41 (0.47-4.20) |
| Vital statistics on Presentation | |
| Calculated PaO2/FiO2 | 0.99 (0.99-1.00) |
| Heart Rate | 1.00 (0.99-1.00) |
| spO2 | 1.03 (1.00-1.06) |
| Respiratory Rate | 1.01 (0.99-1.04) |
| Laboratory Tests on Presentation | |
| D-Dimer | 0.96 (0.93-1.00) |
| Ferritin | 1.00 (0.99-1.01) |
| Lactate Dehydrogenase | 1.00 (0.99-1.01) |
| C-reactive protein | 1.00 (0.99-1.01) |
| Procalcitonin | 0.98 (0.96-1.01) |

COPD denotes chronic obstructive pulmonary disease, ACE angiotensin converting enzyme inhibitors, ARB angiotensin-receptor blockers, DOAC Direct Oral Anticoagulants
*Excluding azithromycin
C-statistic=0.81

Table S3. Sensitivity analyses showing hazard ratio for primary outcome using different baselines and, separately, only considering patients who received treatment prior to baseline time

| Method | HCQ during follow-up and before the composite endpoint or before baseline | | | HCQ before baseline | | |
|---|---|---|---|---|---|---|
| | HCQ (N) | No HCQ (N) | Hazard ratio (95% CI) | HCQ (N) | No HCQ (N) | Hazard ratio (95% CI) |
| 24 hours | | | | | | |
| Crude analysis | 811 | 565 | 2.37 (1.84 - 3.02) | 374 | 1002 | 1.74 (1.40 - 2.16) |
| Multivariable analysis* | 811 | 565 | 1.00 (0.76 - 1.32) | 374 | 1002 | 1.05 (0.83 - 1.33) |
| Propensity Score Analyses | | | | | | |
| IPW** | 811 | 565 | 1.04 (0.82 - 1.32) | 374 | 1002 | 0.93 (0.72 - 1.20) |
| Matching*** | 811 | 274 | 0.98 (0.73 - 1.31) | 374 | 250 | 1.07 (0.80 - 1.42) |
| Adjusting for PS**** | 811 | 565 | 0.97 (0.74 - 1.28) | 374 | 1002 | 1.05 (0.83 - 1.33) |
| 48 hours[#] | | | | | | |
| Crude analysis | 768 | 535 | 3.15 (2.34 - 4.24) | 649 | 654 | 2.01 (1.57 - 2.58) |
| Multivariable analysis* | 768 | 535 | 1.27 (0.91 - 1.77) | 649 | 654 | 0.88 (0.66 - 1.17) |
| Propensity Score Analyses | | | | | | |
| IPW** | 768 | 535 | 1.20 (0.92 - 1.58) | 649 | 654 | 0.81 (0.63 - 1.04) |
| Matching *** | 768 | 234 | 1.27 (0.89 - 1.81) | 649 | 265 | 0.90 (0.67 - 1.20) |
| Adjusting for PS**** | 768 | 535 | 1.22 (0.88 - 1.70) | 649 | 654 | 0.87 (0.66 - 1.15) |

IPW = inverse probability weighting; PS = propensity score; HCQ = Hydroxychloroquine

# Patients are excluded if they reached an endpoint prior to 48 hours; baseline is reset at 48 hours (a landmark analysis)

* Hazard ratio from the multivariable Cox proportional hazards model stratified on sex, chronic lung disease, and body mass index and additionally adjusted for age, race/ethnicity, BMI, insurance, current smoker, past diagnoses, current medications, vital statistics and laboratory tests on presentation

** Hazard ratio from the multivariable Cox proportional hazards with the same strata and covariates with inverse probability weighting by the propensity score

*** Hazard ratio from a multivariable Cox proportional hazards model with the same strata and covariates with matching by the propensity score

**** Hazard ratio from a multivariable Cox proportional hazards model with the same strata and covariates and additionally adjusted for the propensity score

Table S4. Hazard ratios (95% CIs) for the composite endpoint for all variables included as covariates in the Cox multivariable model with inverse probability weighting by the propensity score (primary analysis)*

| | |
|---|---|
| Hydroxychloroquine | 1.04 (0.82-1.32) |
| Age (ref=<40) | |
| 40-59 | 1.52 (0.78-2.93) |
| 60-79 | 2.09 (1.05-4.13) |
| >80 | 3.92 (1.88-8.20) |
| Race/ethnicity (ref=NH White) | |
| NH Black | 0.81 (0.51-1.29) |
| Hispanic | 0.69 (0.47-1.00) |
| Other | 0.53 (0.31-0.90) |
| Insurance (ref=Medicaid) | |
| Medicare | 0.80 (0.56-1.16) |
| No insurance | 0.54 (0.30-0.98) |
| Commercial | 0.57 (0.37-0.87) |
| Current Smoker | 0.79 (0.54-1.16) |
| Past Diagnoses | |
| Essential hypertension | 1.11 (0.84-1.46) |
| Cancer | 0.93 (0.67-1.30) |
| Chronic kidney disease | 0.51 (0.35-0.74) |
| Transplant, HIV, or immune suppressive medications | 0.98 (0.58-1.66) |
| Current Medications | |
| Statins | 1.08 (0.84-1.39) |
| ACE/ARB | 0.86 (0.66-1.11) |
| Steroids | 2.88 (2.17-3.80) |
| DOAC or Coumadin | 0.89 (0.61-1.29) |
| Azithromycin | 1.03 (0.81-1.31) |
| Other Antibiotics | 3.33 (2.25-4.93) |
| Tocilizumab | 1.27 (0.84-1.91) |
| Remdesivir | 1.00 (0.50-2.00) |
| Vital statistics on Presentation | |
| Calculated PaO2/FiO2 | 0.99 (0.99-1.00) |
| Heart Rate | 1.01 (0.99-1.03) |
| spO2 | 1.00 (0.99-1.01) |
| Respiratory Rate | 1.01 (0.99-1.03) |
| Laboratory Tests on Presentation | |
| D-Dimer | 1.01 (0.99-1.03) |
| Ferritin | 1.00003 (1.00000-1.00008) |
| Lactate Dehydrogenase | 1.00084 (1.00045-1.00124) |
| C-reactive protein | 1.00308 (1.00165-1.00451) |
| Procalcitonin | 1.00 (0.99-1.01) |

* This Cox multivariable model, which was inverse probability weighted by the propensity score, was designed to control for potential confounding-by-indication of the exposure (hydroxychloroquine use) and composite outcome. It was additional  stratified on sex, chronic lung disease and BMI, for which parameter estimates are not generated or shown.  The results in this table are provided for the reader's information but should be not be interpreted to provide information on predictors or causes of the composite outcome.

**References**

1. Johnson S, Friedman C, Cimino J, Hripcsak G, Clayton P. Conceptual data model for a central patient database. Proc Annu Symp Comput Appl Med Care 1991:381-5.
2. Hripcsak G, Duke J, Shah N, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers.  MEDINFO. Sao Paulo, Brazil2015.
3. Brown SM, Grissom CK, Moss M, et al. Nonlinear Imputation of Pao2/Fio2 From Spo2/Fio2 Among Patients With Acute Respiratory Distress Syndrome. Chest 2016;150:307-13.
4. Brown SM, Duggal A, Hou PC, et al. Nonlinear Imputation of PaO2/FIO2 From SpO2/FIO2 Among Mechanically Ventilated Patients in the ICU: A Prospective, Observational Study. Crit Care Med 2017;45:1317-24.